Chapter I
Evaluation Framework

According to the *Standards for Conduct and Evaluation of Activities Carried Out by the Office of Educational Research and Improvement (OERI)*, each of the 10 Regional Educational Laboratories (RELs) is to be independently evaluated in the third year of its five-year contract with the Department of Education (ED). While the *Standards* call for each Lab to be evaluated independently rather than as parts of a larger Laboratory system, a standard evaluation framework was developed to guide the peer review process and provide some consistency across sites.

This chapter describes the evaluation framework that was used by the peer review panelists to assess the work of the Laboratories. The framework includes a set of evaluation criteria against which the REL's products, services and activities were assessed through a series of overarching questions. For each of the evaluation questions, a set of evaluation indicators were identified to guide the peer review panelists. The framework also includes a list of possible data sources that could be used to evaluate the various products and services of the Lab. Using this standard framework, DIR developed an individual evaluation plan for each Laboratory, selecting specific activities and data sources to be reviewed that were representative of the work of that particular Lab.

A.   Evaluation Criteria

The evaluation framework is guided broadly by the four evaluation criteria specified in the *Standards*:

- Implementation and Management
- Quality
- Utility
- Outcomes and Impact

The *Standards* provide guidance on how peer reviewers can assess whether a recipient meets the evaluation criteria. Additionally, the *Standards* set forth guidance specific to assessing the performance of the Regional Educational Laboratories. Excerpts from the *Standards* are provided below for each of the evaluation criterion.

Implementation and Management
> *Peer reviewers shall consider the degree to which the recipient has fully executed its program of work. In doing so, peer reviewers shall consider evidence on the extent to which the recipient completes the work described in the approved application or contract, including any approved modifications, in the time period proposed and in an efficient manner.*
>
> *In examining the degree of implementation, peer reviewers may also consider evidence on the extent to which (a) the recipient implements and utilizes a quality assurance system for its products and services or both; and (b) the recipient*

conducts self-assessment or self-evaluation activities, including periodically seeking out independent critiques and evaluations of its work, and uses the results to improve performance.

Quality

Peer reviewers shall consider the degree to which the recipient's work approaches or attains professional excellence. In determining quality, peer reviewers shall consider evidence on the extent to which (a) the recipient utilizes processes, methods, and techniques appropriate to achieve the goals and objectives for the program of work in the approved application; and (b) the recipient applies the appropriate processes, methods, and techniques in a manner consistent with the highest standards of the profession. In determining quality, peer reviewers may also consider the extent to which the recipient conducts a coherent, sustained program of work informed by relevant research.

For the Regional Educational Laboratories, peer reviewers shall also consider evidence on the extent to which: (a) the recipient utilizes a well-conceptualized framework and sound theoretical and methodological tools in conducting professionally rigorous studies ad model meets the following criteria; (b) the recipient conducts work of sufficient size, scope, and duration to provide sound guidance for improvement efforts; and (c) the recipient's products are well tested and based on sound research.

Utility

In determining the utility of the recipient's products or services or both, peer reviewers shall consider evidence on the extent to which the recipient's work (including information, materials, processes, techniques, or activities) is effectively used by and is useful to its customers in appropriate settings. In determining utility, peer reviewers may also consider the extent to which the recipient has received national recognition; e.g., articles in refereed journals and presentations at professional conferences.

For the Regional Educational Laboratories, peer reviewers shall also consider evidence on the extent to which the recipient documents, reports, and disseminates its work in ways to facilitate its effective use in appropriately targeted settings, particularly school improvement efforts of States and localities.

Outcomes and Impact

Peer reviewers shall consider the results of the recipient's work. In examining outcomes and impact, peer reviewers shall consider evidence on the extent to which (a) the recipient meets the needs of its customers; and (b) the recipient's work contributes to the increased knowledge or understanding of educational problems. In examining outcomes and impact, peer reviewers may also consider the extent to which recipients address issues of national significance through its products and services or both.

For the Regional Educational Laboratories, peer reviewers shall also consider evidence on the extent to which: (a) the recipient assists States and localities to implement comprehensive improvement strategies through the provision of research-based information (including well-tested models and strategies), materials, and assistance; and (b) the recipient's work results in widespread access to information regarding research and best practices, particularly within its region.

B.   Evaluation Questions

Eight evaluation questions were developed to correspond to the criteria set forth in the *Standards* and to broadly reflect the work of the Labs. The questions were developed by DIR with input from a Technical Work Group (TWG), the Department of Education's Planning and Evaluation Service (PES), the Office of Educational Reform and Improvement (OERI), and the Labs themselves. The evaluation questions were designed to allow peer reviewers to evaluate the Lab's products and services by addressing the following questions:

- To what extent is the REL doing what it was approved to do during the first three contract years?
- To what extent is the REL using a self-monitoring process to plan and adapt activities in response to feedback and customer needs?
- To what extent is the REL developing high quality products and services?
- To what extent are the products and services provided by the Laboratory useful to, and used by customers?
- To what extent is the REL focused on customer needs?
- To what extent is the REL's work contributing to improved student success, particularly in intensive implementation sites?
- To what extent does the Laboratory assist states and localities to implement comprehensive school improvement strategies?
- To what extent has the REL made progress in establishing a regional and national reputation in its specialty area?

These eight questions are presented again in Appendix I, along with associated evaluation indicators and possible data sources.

C.   Evaluation Indicators

For each of the eight evaluation questions, a set of evaluation indicators was developed to provide guidance to the peer reviewers regarding specific factors to consider in the course of the review process. Several of the indicators relate directly back to language included in the *Standards*. Others were developed by DIR, with input from the TWG, PES, OERI, and the Labs, to reflect important dimensions of the work conducted by the Labs.

In an effort to provide further clarification regarding the evaluation indicators, DIR developed a companion document to the evaluation framework. In this document (Appendix I), a series of questions was provided under each evaluation indicator. The questions suggest

possible dimensions along which peer reviewers might choose to evaluate materials. These dimensions were <u>not</u> meant to be exhaustive but were provided as suggestions of some issues to consider within each indicator. Peer reviewers were in no way restricted to these areas of inquiry and were free to evaluate materials along additional dimensions they felt were appropriate within each indicator. Peer reviewers were not expected to address each of the specific questions included in the companion document in their written reports. Rather, reviewers were asked to address the eight evaluation questions, considering the accompanying indicators as a whole.

D.   Possible Data Sources

The *Standards* specify several data sources that a peer review panel must consider for an interim assessment. These include:

- *The original request for proposals or grant announcement and the contract proposal or grant application*
- *Documentation of any changes in the work described in the contract, grant, or cooperative agreement, including reasons for changes*
- *Any progress reports delivered to the Department or made available to the public by the recipient*
- *Examples of products delivered to the Department or made available to the public by the recipient*
- *Any relevant reports written by OERI staff, including reports of site visits made by OERI staff*
- *Any performance evaluations conducted under the FAR or the Education Department General Administrative Regulations (34 CFR Part 75)*
- *Any relevant information provided by the recipient in response to the Government Performance and Results Act (GPRA) requirements*
- *Any reports from program evaluations commissioned by the Department*

In addition to these mandated data sources, the *Standards* suggest that self-assessments, site visits, oral or written presentations, and other information about the recipient's performance may also be considered in the interim assessment.

While the Interim Evaluation was designed to assess Laboratory performance in the first three years of the current contract period, it was neither reasonable nor feasible for panelists to review all of the work conducted to date. Instead, peer reviewers were asked to address the evaluation questions based on a review of selected data sources sampled from each of the following categories: (1) "signature works," (2) additional selected outputs, and (3) Laboratory operations materials. Possible data sources included printed and electronic materials, in addition to on-site interviews and observations. The list of possible data sources presented is not meant to be exhaustive, but is representative of the types of materials and information that were available to reviewers.

<u>Signature Works</u>. For the purpose of this evaluation, a "signature work" was defined as a significant illustrative work (e.g., major product, strategic thrust, or service package) that best exemplifies the work of the Laboratory in the current contract period. While it was understood

that the signature works may not represent a typical cross section of the work conducted at a Laboratory, their inclusion was meant to provide peer reviewers with a more in-depth look at some of the major program areas of the Laboratory. Labs were asked to nominate up to six signature works, of which two would be selected for inclusion in the Interim Evaluation. (A more detailed discussion of the selection of signature works is provided in Chapter III). The concept of a signature work was not meant to minimize the importance of other work conducted at the Laboratory, but rather to focus the review process in a way that was manageable for the reviewers to complete given the limited time and resources available.

To be considered a signature work, the nominated work had to meet the following criteria:

- Be central to the major strands of work as described in the REL technical proposal, modifications, or annual update
- Represent a significant percentage of REL allocated resources (in $ and/or staff)
- Be primarily funded by REL monies and/or represent a significant portion of the funding within a particular REL task
- Have the potential for "scaling up" and expansion
- Be work accomplished in this contract period (December 1995-December 1998)
- Be supported by sufficient data sources to facilitate a thorough review of the work

For the purposes of this Interim Evaluation, DIR was specifically interested in signature works where the bulk of the work had been conducted in this contract period. It was understood that some projects or programs may have had an initial phase or start-up work in a previous contract period, and/or have plans to continue another phase or related piece into the next contract period. However, only those portions of a program/project/product funded by the REL contract and accomplished within this contract period were eligible for review.

Additional Selected Outputs. In addition to the selected signature works, peer reviewers were provided with a sample of other programs, products, and services produced by the Lab. This review of other outputs was intended to provide peer reviewers with a broader base upon which to gauge the totality of activities of the Lab under the OERI/REL contract. To help identify the universe of possible data sources in this category," each Laboratory was asked to prepare an inventory of the major tangible products developed or activities/services conducted during the first three years of the contract period (i.e., December, 1995-December, 1998). The inventory was not designed to be a log of every product, service, and/or program developed or provided in the contract period, but was intended to be complete enough to capture the depth and range of ongoing work of the Laboratory.

Laboratory Operations Materials. The *Standards* mandate a series of data sources that address how the Lab operates at the organizational level. In addition to the required items (e.g., original contract proposal, contract modifications, quarterly reports, annual plans and updates), Labs were asked to identify other materials or data that would allow the peer reviewers to address questions of implementation and management. These other data sources often included organizational charts, Board rosters and minutes, and REL funding and staff allocations by task.